The EMBO Journal vol.12 no.2 pp.513-525, 1993

Recognition of the *Xenopus* ribosomal core promoter by the transcription factor xUBF involves multiple HMG box domains and leads to an xUBF interdomain interaction

Benoît Leblanc, Chris Read and Tom Moss[1]

Centre de Recherche en Cancérologie de l'Université Laval, Hôtel-Dieu de Québec, 11 Côte du Palais, G1R 2J6 Québec, Canada

Communicated by I.Grummt

[1]Corresponding author

The interaction of the ribosomal transcription factor xUBF with the RNA polymerase I core promoter of *Xenopus laevis* has been studied both at the DNA and protein levels. It is shown that a single xUBF–DNA complex forms over the 40S initiation site (+1) and involves at least the DNA sequences between −20 and +60 bp. DNA sequences upstream of +10 and downstream of +18 are each sufficient to direct complex formation independently. HMG box 1 of xUBF independently recognizes the sequences −20 to −1 and +1 to +22 and the addition of the N-terminal dimerization domain to HMG box 1 stabilizes its interaction with these sequences ~10-fold. HMG boxes 2/3 interact with the DNA downstream of +22 and can independently position xUBF across the initiation site. The C-terminal segment of xUBF, HMG boxes 4, 5 or the acidic domain, directly or indirectly interact with HMG box 1, making the core promoter sequences between −11 and −15 hypersensitive to DNase. This interaction also requires the DNA sequences between +17 and +32, i.e. the HMG box 2/3 binding site. The data suggest extensive folding of the core promoter within the xUBF complex.

Key words: ribosomal/RNA polymerase I/transcription factor/xUBF/*Xenopus laevis*

Introduction

In eukaryotes, ribosomal transcription is undertaken in the nucleolus by a dedicated RNA polymerase, RNA polymerase I (Pol I). Apart from the multipeptide Pol I complex, whose activity appears to be growth regulated (Bateman and Paule, 1986; Tower and Sollner-Webb, 1987), ribosomal transcription has been shown to require three other protein factors; UBF and a second factor, variously called SL-1, TIF-IB, TFID or Factor D (Bell *et al.*, 1988; Schnapp *et al.*, 1990a; Smith *et al.*, 1990; Tanaka *et al.*, 1990; McStay *et al.*, 1991b; Paule *et al.*, 1991), which interact with the Pol I specific promoter sequence and a third factor, variously called TIF-IA, Factor C or TFIC (Mahajan and Thompson, 1990; Schnapp *et al.*, 1990b), which probably interacts with the polymerase. The activities of both UBF and the polymerase associated factor are also growth regulated (Bateman and Paule, 1986; Gokal *et al.*, 1990; Mahajan *et al.*, 1990; Schnapp *et al.*, 1990b; Voit *et al.*, 1992) and the expression of UBF is regulated by differential splicing in mouse (Hisatake *et al.*, 1991) and during early development and differentiation in *Xenopus* (Guimond and Moss, 1992). SL-1 was shown to be a complex of three or four peptides, one of which was the TATA-box binding protein (TBP), which is now known to be implicated in transcription by all three RNA polymerases (Comai *et al.*, 1992; Cormack and Struhl, 1992; Schultz *et al.*, 1992; White *et al.*, 1992). UBF and the SL-1 complex interact in a cooperative manner with the ribosomal promoter to form a stable promoter complex with which active Pol I and its associated factor specifically interact. Hence promoter selectivity is a property of UBF and SL-1 (see Sollner-Webb and Mougey, 1991). Of these two factors, UBF appears to be able independently to form a complex with the promoter and to provide a binding site for SL-1. Hence the formation of a Pol I specific complex at a ribosomal promoter probably begins with the binding of UBF.

UBF was initially characterized in human (hUBF) by its interaction with the upstream control element of the ribosomal promoter (Bell *et al.*, 1988). In *Xenopus* it was also shown to interact with the tandemly repeated enhancers (Pikaard *et al.*, 1989). Subsequently, the hUBF and then the xUBF cDNAs were cloned, revealing a very unusual protein of between 80 and 90 kDa, the first of a family of HMG box transcription factors (Jantzen *et al.*, 1990; Bachvarov and Moss, 1991; Bachvarov *et al.*, 1991). The HMG box is a DNA binding motif originally found in HMG 1 and 2 (Reeck *et al.*, 1982; Cary *et al.*, 1983). *Xenopus* UBF (xUBF) contains five tandemly repeated HMG box homologies followed at the C-terminus by two blocks of solely acidic residues (Bachvarov and Moss, 1991). Other HMG box factors include the sex determination factor SRY, the lymphoid and T cell specific LEF-1 (TCF1α) and TCF1, the mitochondrial factors mtTF1 and ABF2 (Gubbay *et al.*, 1990; Kolodrubetz and Burgum, 1990; Sinclair *et al.*, 1990; Diffley and Stillman, 1991; Parisi and Clayton, 1991; Travis *et al.*, 1991; van de Wetering *et al.*, 1991). The HMG box of LEF-1 has been shown to contact mainly the minor DNA groove and both LEF-1 and SRY have been shown to bend the DNA duplex considerably (Giese *et al.*, 1991, 1992). The HMG boxes of HMG 1 and 2 have not as yet been shown to display any DNA sequence recognition, except for a slight preference for dA and dT rich DNA, but they do bind strongly to negatively supercoiled DNA and especially to potential cruciform structures (see Lilley, 1992). Hence the HMG boxes may also bind preferentially to bent DNA. In contrast to HMG 1 and 2, xUBF shows clear, though as yet ill defined, DNA sequence preferences and *Xenopus* and rat UBFs dimerize via an N-terminal domain which has been shown to be essential for efficient *in vitro* transcription (McStay *et al.*, 1991a; O'Mahony *et al.*, 1992). However, in common with HMG1 and 2, xUBF has an essentially purely acidic C-terminal domain. Such domains have been shown to be most probably functionally distinct from the 'acidic activator' domains of some polymerase II transcription factors (Landsman and Bustin, 1991) but in the

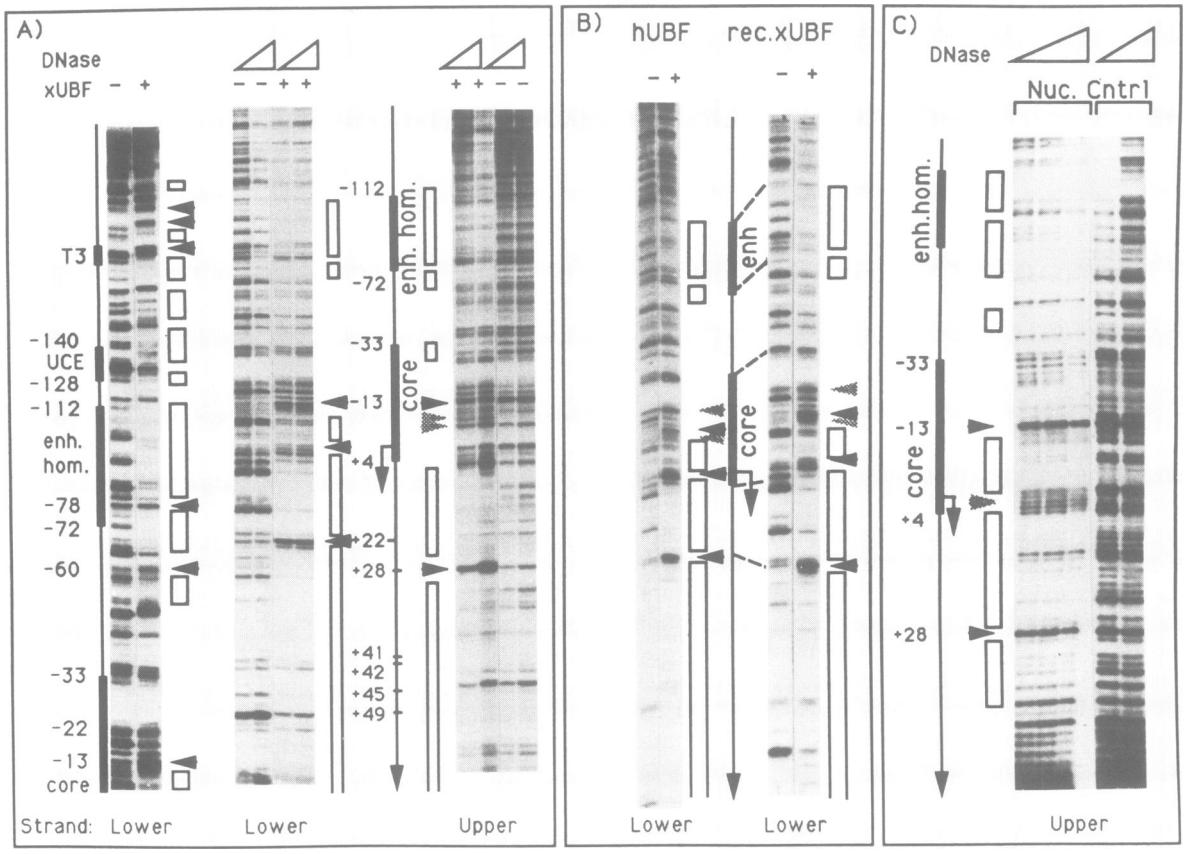

Fig. 1. xUBF produces a characteristic DNase I protection of the *X.laevis* core promoter and 5′ transcribed sequences. (**A**) The *in vitro* DNase I footprints of tissue culture purified xUBF in the region of the *X.laevis* 40S promoter. (**B**) Footprints of recombinant human UBF (hUBF) and recombinant xUBF (rec.xUBF) on the lower strand of the 40S core promoter and flanking sequences. (**C**) *In vivo* DNase I footprinting of the 40S core promoter and flanking sequences. Footprints performed on isolated nuclei are shown (Nuc.) alongside controls performed on purified DNA (Cntrl). In comparing the *in vitro* and *in vivo* data it should be noted that a certain degree of sequence heterogeneity (Stewart *et al.*, 1983) exists within the 450 haploid gene copies. 'core', 'enh. hom.', 'UCE' and 'T3′ respectively refer to the core promoter, enhancer homology, upstream control region and promoter adjacent terminator. Numbering is given from the 40S transcription initiation site at +1 and the direction of transcription is indicated with a vertical arrow. Protected regions are indicated by empty boxes and DNase I hypersensitive sites by arrows.

case of xUBF and mouse UBF are important for efficient *in vitro* transcription by Pol I (Voit *et al.*, 1992; McStay *et al.*, 1991a).

Here we investigate promoter recognition by xUBF and the role of the HMG boxes in this process. We show that specific recognition of the core promoter depends on multiple partially redundant interactions spread over an 80 bp region and leads to a folding of the initiation site sequences.

Results

Studies on the interaction of the transcription factor UBF with the ribosomal genes have concentrated on the enhancers and the upstream control element (UCE) of the promoter. However, it is evident from the available data on the mammalian polymerase I promoter, that UBF must also functionally interact with the so-called core promoter. In fact in *Xenopus* one of the most extensive and well defined regions of xUBF interaction occurs within the core promoter and rather surprisingly continues well downstream of the initiation site (Read *et al.*, 1992).

Figure 1A summarizes the DNase I footprints produced by xUBF, in this case purified from tissue culture cells, on the *X.laevis* ribosomal (40S) promoter. Footprinting within the promoter in the region of the UCE and the enhancer homology was clearly evident as well as further upstream around the terminator (T3) and in the proximal enhancer unit. xUBF interactions with the core promoter were characterized by a series of hypersensitive cleavages (hypersites) around −20, −13 and −1. However, one of the most striking xUBF interactions was found to occur downstream of the transcription initiation site (+1). Figure 1A shows two extensive regions of protection on both DNA strands starting immediately downstream of +1 and separated by a hypersensitive cleavage at +22 on the lower and at +28 on the upper strand, i.e. on the same face of the DNA duplex. This protection, which we will hereafter refer to simply as the 'downstream' footprint was due solely to xUBF binding, since both recombinant vaccinia expressed human UBF and bacterially expressed xUBF gave an identical protection (Figure 1B).

In vivo footprinting also revealed a very analogous pattern of protection and hypersensitivity upstream and downstream of +1 on the ribosomal genes of mature erythrocytes (Figure 1C). On the upper DNA strand, two regions of protection immediately downstream of the initiation site and separated by a hypersite were evident. The position of this hypersite was mapped to +28, i.e. identical to that seen with xUBF *in vitro* (Figure 1A). The *in vivo* footprint also indicated other similarities with xUBF binding *in vitro*, the initiation site and the −13 region being accessible to DNase.

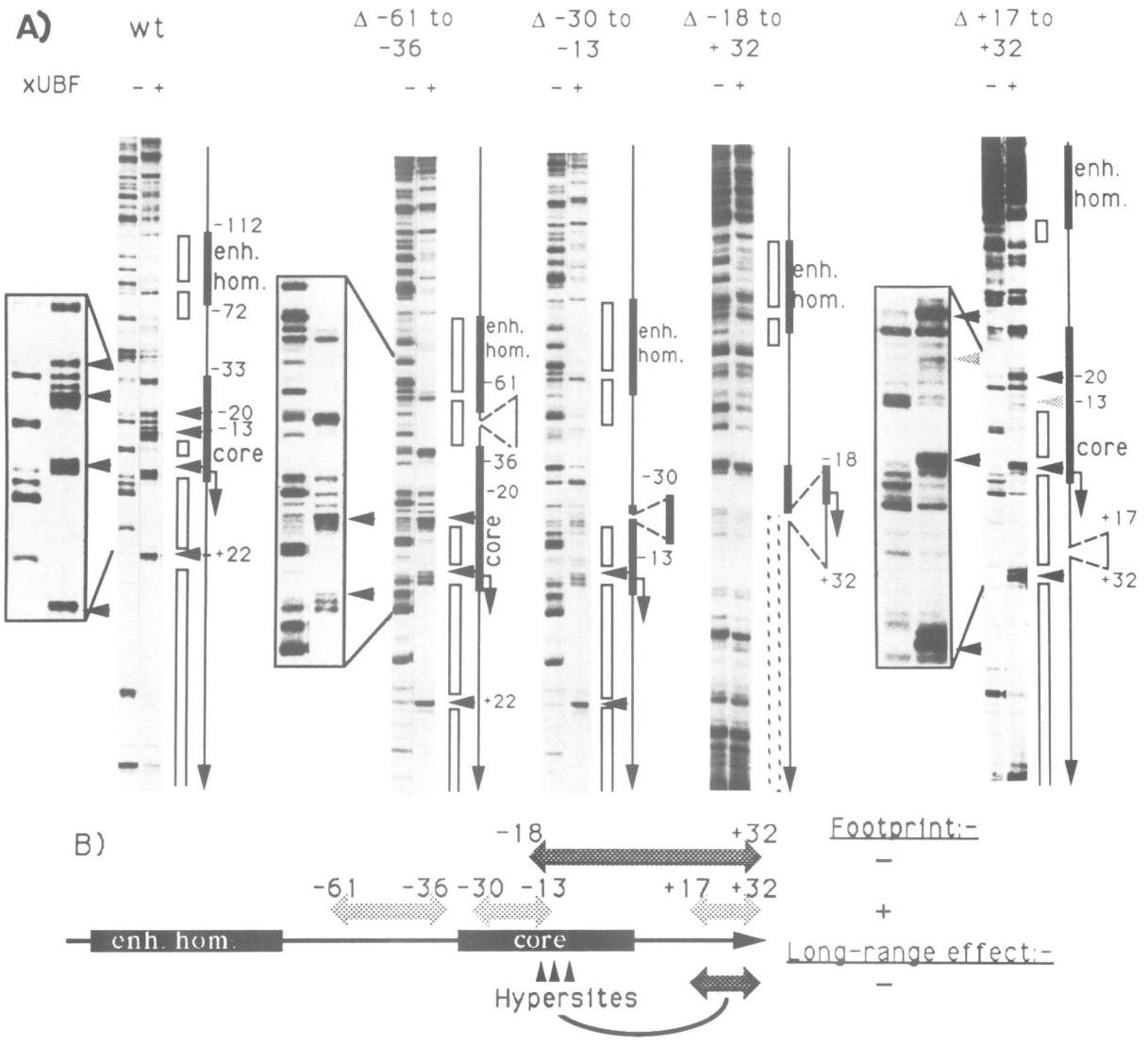

Fig. 2. Sequences essential for xUBF positioning map around the 40S initiation site. (**A**) xUBF from tissue culture was DNase I footprinted on the lower strand of the wild-type promoter (wt) and promoter deletion mutants (Δ-). The dashed box refers to the expected position of footprinting on the Δ −18 to +32 mutant. (**B**) Summary of the internal deletion data. The extents of the deletions are indicated by arrows. 'Long-range effect' refers to the DNA sequences required to produce hypersensitivity (Hypersites) between −11 and −15 (see text). Otherwise, the nomenclature is as in Figure 1.

A similar pattern of protection could be discerned, albeit less clearly, on the ribosomal promoters of rapidly dividing *Xenopus* tissue culture cells (manuscript in preparation).

The data of Figure 1C were consistent with xUBF binding downstream of the 40S initiation site *in vivo* and hence of this interaction playing some role *in vivo*. Considering that *X.laevis* has 450 copies of the ribosomal genes per haploid genome, the observation of such footprinting and hypersensitivity also suggested a very high degree of homogeneity in gene organization *in vivo*.

The downstream xUBF interaction is directed by sequences around the 40S initiation site

The finding that an essential transcription factor interacts both *in vitro* and probably *in vivo* with sequences downstream of the 40S initiation site was rather surprising, since the *X.laevis* promoter has been mapped upstream of about +4 (Read *et al.*, 1992 and references therein). However, xUBF clearly interacts with a bewildering range of apparently unrelated DNA sequences, e.g. see Figure 1A. In fact to date it has not been possible to define a convincing consensus binding sequence nor even to define the length of DNA with which a single xUBF molecule or xUBF dimer interacts. Given that xUBF is one of only two essential transcription factors which interact directly and selectively with the ribosomal promoter, its binding immediately downstream of +1 seemed unlikely to be fortuitous. Mammalian promoters have indeed been shown to extend as far downstream as +16 or +20 (see Moss *et al.*, 1985). The potential for multiple interactions via the repeated DNA binding domains within one molecule also makes interaction of the UBFs with DNA a fascinating study in itself. We therefore used the downstream xUBF footprint as a model interaction by which to resolve some of these questions.

By footprinting xUBF on deletion mutants of the 40S promoter it was shown that large segments of both the upstream and downstream promoter regions could be removed without affecting the downstream footprint

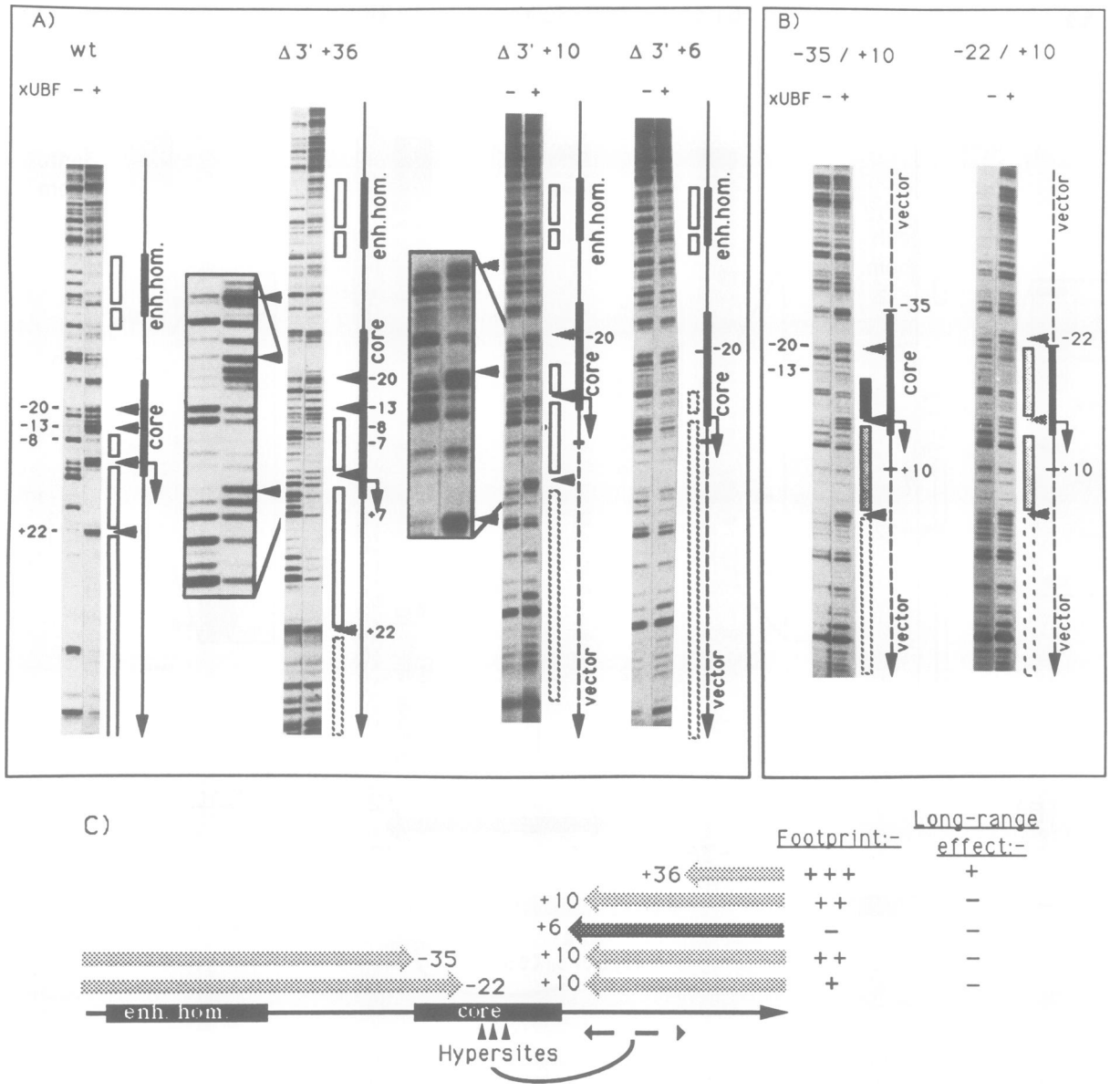

Fig. 3. 3′ and 5′ boundaries of the xUBF recognition sequence in the core promoter. xUBF from tissue culture was footprinted on the lower strand of (**A**) the wild-type promoter (wt) and promoter constructs deleted on the 3′ side (Δ3′-) or (**B**) promoter constructs deleted on both the 5′ and 3′ sides. Dashed lines refer to non-ribosomal, plasmid vector sequences (vector). Open boxes indicate regions of protection, shadowed boxes of partial protection and dashed boxes, regions of expected protection. (**C**) Summary of the 3′ and 5′ deletion data. The extents of the deletions are indicated by arrows. 'Long-range effect' refers to the presence of DNase hypersensitivity (Hypersites) between −11 and −15 (see text). Otherwise, the nomenclature is as in Figure 1.

(Figure 2A, see 2B for summary). Deletions of −61 to −36, −30 to −13 and even +17 to +32 left the position of the downstream footprint and hypersensitive sites unchanged, despite the fact that in the latter case the hypersensitive site now fell in an unrelated sequence. In each case the enhancer homology footprint was also observed. Deletion of −18 to +32 did, however, eliminate the downstream footprint without affecting the footprint in the enhancer homology. Thus, these two footprinting regions represented independent xUBF interactions. From the −30 to −13, −18 to +32 and +17 to +32 deletions it was tentatively concluded that the downstream footprint required an interaction of xUBF within −12 to +16.

Deletion of all sequences downstream of +36 (Δ3′+36) and their replacement with vector sequence had little effect on xUBF protection between +1 and the hypersensitive site at +22 (Figure 3A and C) though this site and protection further downstream became somewhat less apparent. After extension of the 3′ deletion up to +10, the +22 hypersite was still clearly observed, though it now occurred in the adjacent vector sequences. Protection of the sequences immediately downstream of +1 was also still observed. However, extending the 3′ deletion up to +6 essentially eliminated all signs of the downstream interaction. From these data it was concluded that sequences upstream of +10 were sufficient to position xUBF correctly downstream of the initiation site. Hence, this was in complete agreement with the internal deletion data of Figure 2.

Support for the notion that the core promoter and the downstream region formed a single xUBF binding site came from comparison of the 3′+10 and +6 deletions (Figure 3). The +22, +1, −1 and −20 to −22 bases were found to

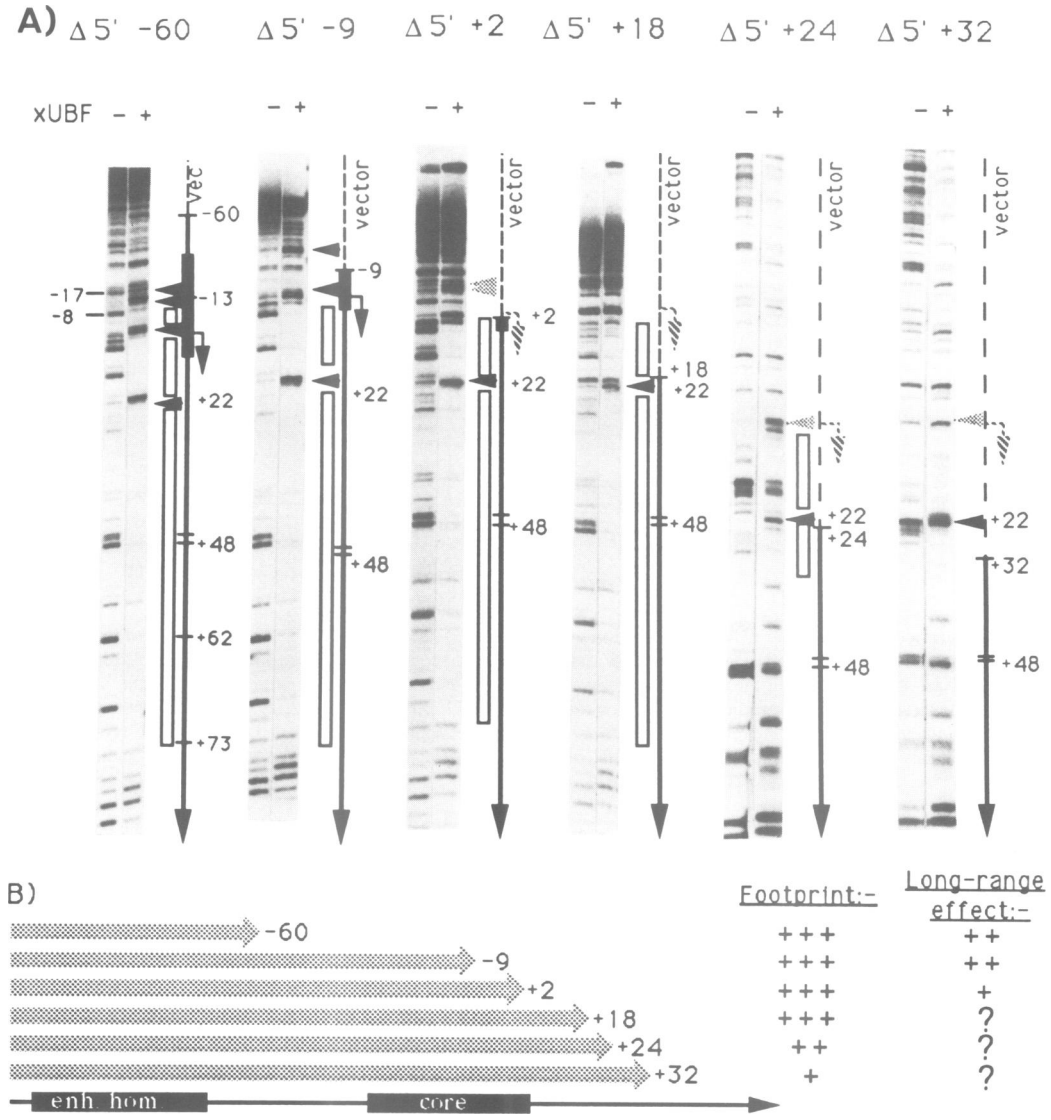

Fig. 4. Correct positioning of xUBF around the initiation site can also be directed by sequences outside of the mapped promoter. Dashed lines refer to non-ribosomal, plasmid vector sequences (vector). (**B**) Summary of the 5′ deletion data. The extents of the deletions are indicated by arrows. 'Long-range effect' refers to the presence of hypersites between −11 and −15 (see text). In the cases indicated by '?' the data were inconclusive on this point. Otherwise, the nomenclature is as in Figure 1.

be hypersensitive in the wild type situation and remained so on the 3′+10 deletion. However, on the 3′+6 deletion essentially all signs of this hypersensitivity disappeared along with the disappearance of signs of the downstream interaction. Interaction of xUBF with the enhancer homology remains unaffected in the Δ3′ deletion mutants and hence acted as an internal control.

Thus it appeared that the region from −22 at least to the hypersite at +22 may form a single cooperative xUBF binding site.

A minimal recognition site for xUBF

In order to define the minimal DNA upon which xUBF could position itself correctly, all ribosomal sequences downstream of +10 and upstream of −35 or −22 were deleted (Figure 3B and C). The −35 to +10 sequence directed correct positioning of xUBF, the hypersites around −20, +1 and +22 all being clearly formed, though protection of the intervening regions was notably weaker than on the wild type template. The same characteristics could still be discerned on the −22 to +10 template. Thus a core sequence between −22 and +10 was sufficient to position xUBF correctly, but the quality of the protection achieved suggested that the interaction had been considerably weakened.

Sequences downstream of +10 affect the interaction of xUBF with the core promoter

The removal of sequences downstream of +10 and their replacement with vector sequences did to some extent affect xUBF binding. The footprints were shorter on the downstream side, not significantly overlapping into the vector sequences beyond the hypersite at −22 (Figure 3). The most striking change however, occurred not adjacent to these deletions, but upstream within the core promoter. On the wild type promoter, binding of xUBF led to protection around −8 (and partially of −17) and hypersensitive cutting at +1, −1, −11 to −15 and −20 to −22 (Figure 3A) and 'long-range effect' in Figure 3C. Unexpectedly the 3′ deletion to +10 eliminated one of these features, the hypersensitivity of −11 to −15, while on the 3′+36 deletion

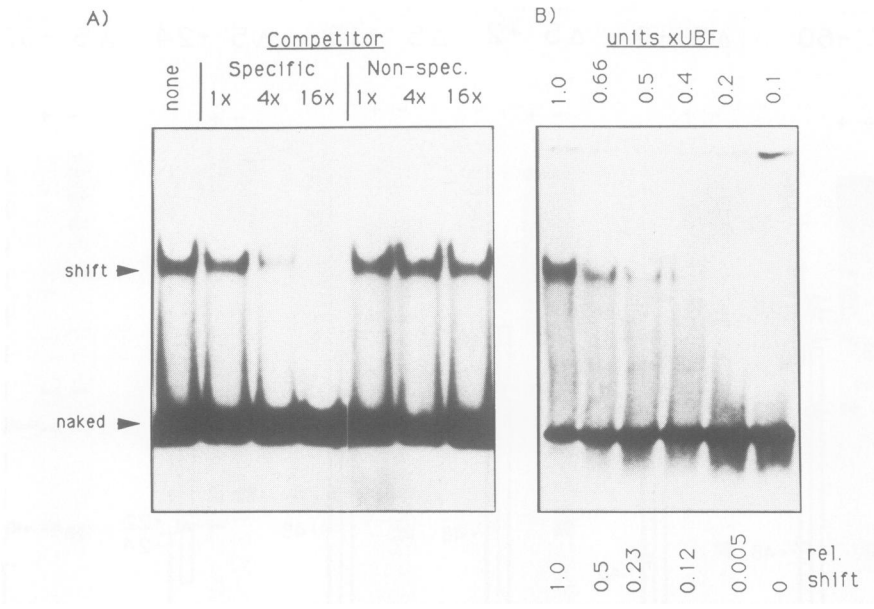

Fig. 5. Gel-shifting demonstrates that a unique xUBF complex forms over the 40S initiation site. A DNA fragment of 155 bp, spanning the ribosomal promoter sequences −9 to +92 and extended on its 5′ side with vector sequence, was used in an xUBF gel-shift assay (Materials and methods). (**A**) The xUBF gel-shift is shown in the absence of competing DNA (none) and competed with a 1- to 16-fold molar excess of either the unlabelled ribosomal promoter fragment, as specific competitor or a pBR322 *DdeI*−*Hin*dIII fragment, as non-specific competitor (non-spec.). (**B**) Titration of the gel-shift with increasing xUBF protein purified from tissue culture. Units are arbitrary, referring to the relative amount of the protein preparation used (see Materials and methods). The relative amount of shifted DNA, as determined by densitometry, is indicated below each lane (rel. shift).

this hypersensitivity was evident. This further confirmed that the core promoter and the downstream sequences formed a single xUBF interaction site. It also showed that interaction of xUBF with the sequences between +36 and +10 had a long-range effect on xUBF interaction with the core promoter (Figure 3C). Referring back to Figure 2 it can be seen that the same repression of hypersensitive cleavage of the bases −11 to −15 was also noted on the +17 to +32 deletion, allowing the sequences required for this long-range effect to be more precisely delineated.

It was concluded that interaction of xUBF with sequences between +17 and +32 induced a long-range effect on xUBF binding around −13 within the core promoter.

Sequences downstream of +18 can also direct the downstream xUBF interaction

The data in Figures 2 and 3 had indicated that sequences downstream of +10 were dispensable for xUBF positioning around the 40S initiation site, but that they did play a role in generating a long-range interaction within the xUBF−DNA complex. To determine whether the downstream sequences could also independently position xUBF, a series of 5′ template deletions was studied (Figure 4). As expected, deletion of all upstream ribosomal sequences (5′ deletion) down to −60 (Δ5′−60), did not alter the characteristics of the xUBF−core promoter interaction. 5′ deletion to −9 and +2 did not affect the formation of the two regions of extensive protection from +1 to +22 and +22 onwards and the hypersites around +1 and +22. Further, the 5′−9 deletion still showed clear hypersensitivity around −15 ('long-range effect' in Figure 4B) and the 5′+2 deletion hypersensitivity around −15 and +1, despite the fact that these sequences had now been replaced by vector DNA. 5′ deletion to +18 left the downstream protection and +22 hypersensitivity unaffected and even 5′ deletion to +24 still gave hypersensitivity around +1 and +22, both sites now falling in vector DNA. Yet further 5′ deletion to +32 did not completely succeed in eliminating the downstream xUBF interaction, the hypersites around +1 and +22 still being visible.

It was concluded that sequences downstream of +18 and possibly even +24, were sufficient to position xUBF correctly downstream of +1 and most probably also within the core promoter. Hence non-overlapping sequences, those between −35 and +10 (Figures 2B and 3C) and those downstream of +18 (+24) (Figure 4B) were each able to position xUBF correctly across the 40S initiation site. In general the more extensive deletions studied gave rise to poorer xUBF protection of the template, e.g. Δ3′+10 and −35 or −22 to +10 in Figure 3 and Δ5′+24 and +32 in Figure 4. This could now be explained in terms of cooperative interactions with non-overlapping sequences within a single xUBF−DNA complex.

Gel shifting demonstrates that a single cooperative xUBF−DNA complex forms at the 40S initiation site

The deletion studies strongly suggested that the xUBF interactions with the core promoter and downstream region represented a single cooperatively binding complex. Mutants which eliminated the downstream interaction also eliminated the characteristic core promoter protection. Further, sequences between +17 and +32 were clearly required to induce hypersensitivity within the core promoter.

If a single cooperative xUBF interaction occurred it should give rise to only one gel-shifted molecular species. Figure 5 shows that a fragment which included only the ribosomal sequences from +9 to +91 (Δ5′+9) gave a single gel-shift. The shift was resistant to competition with a large excess of pdA-T carrier and a non-ribosomal DNA fragment, but was easily competed by unlabelled template (Figure 5A).

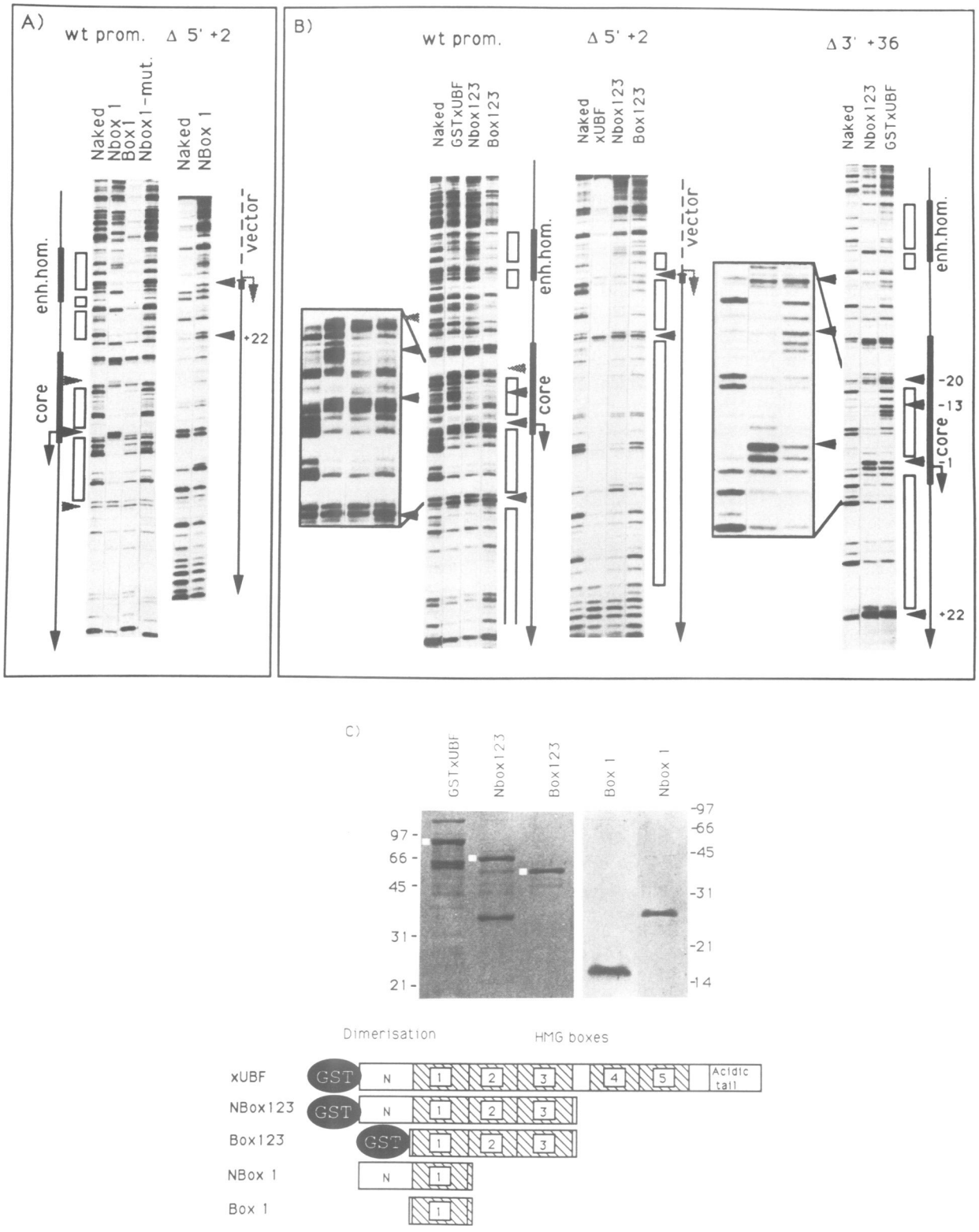

Fig. 6. Correlation between the protected promoter regions and the Hμg boxes of xUBF. (**A**) The NBox1 (0.5 μg) and Box1 (0.9 μg) (C below) and NBox1-mut. (0.4 μg) (see Figure 7), constructs were footprinted on the lower strand of the wild-type 40S promoter and NBox1 (1 μg) also on the lower strand of the Δ5′+2 deletion mutant. (**B**) The xUBF (0.5 μg), Nbox123 (0.6 μg) and Box123 (0.6 μg) GST fusion proteins were footprinted on the wild type promoter, Nbox123 (3 μg) and Box123 (4 μg) on the Δ5′+2 mutant and Nbox123 (3 μg) on the Δ3′+36 mutant. The xUBF track on Δ5′+2 used tissue culture purified xUBF and that on Δ3′+36 used 5 μg of the xUBF–GST fusion protein. It should be noted that in a separate experiment (not shown) the GST domain on Nbox123 and Box123 was removed with thrombin. That did not affect the footprints observed. Non-ribosomal vector sequence is indicated as a dashed line. Otherwise the nomenclature in (A) and (B) is as in Figure 1. (**C**) The xUBF deletion constructions. SDS gel analysis of the purified *E.coli* expressed xUBF deletion constructs is shown above a diagrammatic representation of their structures.

The shifted band decreased in a non-linear manner with decreasing input protein, suggesting a concentration dependent association (Figure 5B). Neither smaller nor larger specific complexes were detected within the range of xUBF concentrations studied. The gel-shift was also detected with highly purified recombinant xUBF (data not shown). It should be noted that this is the first report of a promoter specific gel-shift assay for xUBF.

Together the footprinting and gel-shift data demonstrated conclusively that xUBF formed one single complex with the core promoter and downstream region.

HMG box 1 binds immediately upstream and downstream of the 40S initiation site

xUBF contains five HMG boxes, each potential DNA binding domains, and an N-terminal dimerization motif (Figure 6C). We wished to determine which of these domains were involved in the DNA interaction of xUBF and to which DNA sequences they bound. The HMG box from the transcription factor LEF-1 was shown to bind to its recognition sequence as a monomer and to protect a region of 18−20 bp (Giese *et al.*, 1991). This is consistent with both the expected globular dimensions and DNA bending characteristics of the HMG box (Giese *et al.*, 1992). Thus the very extensive downstream protection due to xUBF suggests the interaction of multiple HMG boxes with the DNA.

As was demonstrated in Figure 1B, recombinant xUBF interacted with the 40S promoter in a manner indistinguishable from the wild type protein. Hence a series of xUBF deletion mutants was expressed and purified from *E.coli* and their DNA interactions studied (Figure 6). A mutant containing the N-terminal dimerization domain and HMG box 1 (Nbox1) efficiently protected the regions from the hypersite at +22 to that at −20/−22 as well as creating a very clear hypersite at +1/−1 and significant hypersensitivity at −20/−22 (Figure 6A). In repeated experiments no significant protection of sequences downstream of +22 was noted, but the Nbox1 mutant did footprint in the region of the enhancer homology.

Removal of the N-terminal domain to leave only HMG box 1 (Box1) caused a very drastic reduction in DNA affinity (see below). This was consistent with the role of the N-terminal domain in dimerization (McStay *et al.*, 1991a; O'Mahony *et al.*, 1992) and strongly suggested that the Nbox1 protein bound around the 40S initiation site as a dimer. At elevated protein concentrations the Box1 mutant was still able to selectively protect the core promoter upstream of −1 as far as −20 and downstream as far as +22, though all hypersites were very poorly developed (Figure 6A). Thus even without the possible cooperativity due to dimerization, both sites −1 to −20 and +1 to +22 were still recognized.

Since we expected a single HMG box to protect ~20 bases (Giese *et al.*, 1991), these data strongly suggested that Nbox1 bound as a dimer, the HMG box of one monomer binding within the core promoter from about −1 to −20 and that of the other monomer binding between +1 and +22. The dimerization of the monomers via their N-terminal domain allowed cooperative binding to the two sites. In order to test this, the footprinting of the Nbox1 protein on promoter deletion mutants was also studied (Figure 6A). It was argued that if one of the two HMG box1 binding sequences was eliminated, binding of Nbox1 should be affected. Footprinting on the 5′+2 mutant was clearly less efficient than on the wild type promoter and no protection downstream of +1 was detected (see Figure 4 for the xUBF footprint on this mutant). However, hypersites around +1 and +22 were still apparent, indicating that some binding did occur, but was presumably less stable allowing DNase attack in normally protected regions.

From these data it was concluded that Nbox1 bound across the 40S initiation site as a dimer, the two HMG box 1 domains interacting with the sites −20 to −1 and +1 to +22. This would predict a palindromic disposition of recognition sequences for HMG box1 around the 40S initiation site (see Discussion).

The interaction of HMG boxes 2 and 3

In order to study the relative importance of the other HMG boxes in xUBF−DNA interaction, two further xUBF constructs were expressed and purified from *E.coli*. Nbox123 contained the N-terminal dimerization domain along with HMG boxes 1−3, while Box123 lacked the dimerization domain (Figure 6C). Both these constructs gave essentially wild type footprints on the intact 40S promoter (Figure 6B). Hypersites at −20 to −22, −1 and +1, and +22 were similar in intensity to the wild type control and clear protection between these sites was evident, though somewhat less so with Box123 than Nbox123. Nbox123 and to some extent Box123, also reproducibly gave nearly wild type levels of protection downstream of the hypersite at +22 as far as about +60 (see also footprint on Δ5′+2). Since the Nbox1 protein did not protect this region of the DNA at all (Figure 6A) it was concluded that HMG boxes 2 and/or 3 must interact downstream of +22. The very extended footprint of xUBF downstream of +22 as far as about +60 could be explained if both boxes contacted DNA and each protected ~20 bp of DNA.

In contrast to the Box1 construct, the N-terminal dimerization domain appeared to be in major part dispensable for binding when HMG boxes 1, 2 and 3 were present (see wild type promoter in Figure 6B). Consistent with this, removal of sequences upstream of +2 had little effect on Nbox 123 binding (see Δ5′+2 in Figure 6B). This further confirmed that HMG boxes 2 and/or 3 must play a significant role in DNA binding.

The DNA affinity of xUBF derives both from dimerization and multiple HMG box binding

The amount of wild type and mutant xUBF required to obtain the core-promoter and downstream footprint was determined by titrating given amounts of wild type promoter fragment with increasing quantities of the xUBF proteins. The dissociation constant of each construct was then estimated from the degree of DNA protection/hypersensitivity it afforded to the initiation site footprint through the range of protein concentrations (see Materials and methods). The DNA affinities of the different xUBF deletion constructions clearly revealed that both HMG boxes 2 and 3 and the N-terminal domain played significant roles in the DNA affinity of xUBF. xUBF purified from *E.coli* gave a K_d of ~14 nM, Nbox123 ~20 nM, Box123 ~120 nM, Nbox1 ~100 nM and Box1 ~1.2 μM. Clearly Box1 bound nearly a 100-fold less strongly than xUBF. Nbox123 had a similar affinity to xUBF, while NBox1 and Box123 both had 5−10 times lower affinities than the full length protein. Thus, both the N-terminal dimerization domain and HMG boxes 1, 2

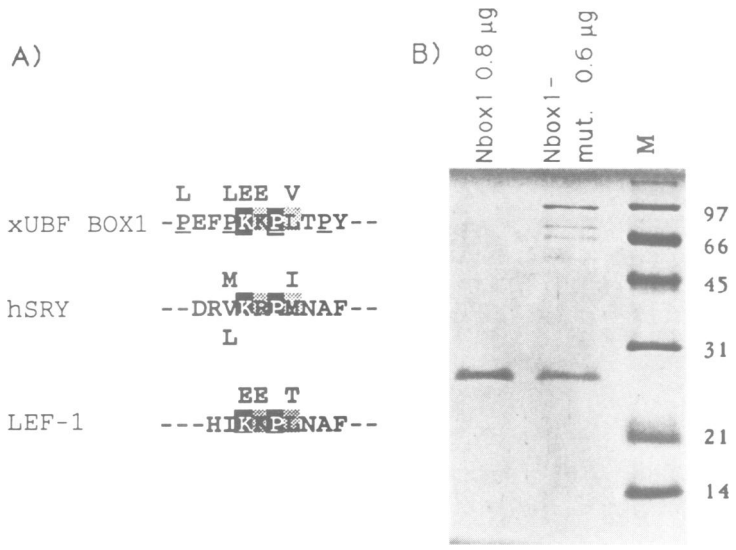

Fig. 7. The HMG box proline repeat is important for DNA binding. (**A**) The N-terminal HMG box regions of the transcription factors SRY and LEF-1 are shown aligned with the N-terminal sequence of xUBF HMG box1 (xUBF BOX1). Mutations which affect DNA binding of SRY and LEF-1 are shown above and below the respective sequences. Mutations introduced into the Nbox1 construct (Figure 6C) are also shown above the xUBF HMG box1 sequence. Conserved and semi-conserved amino acids are shown highlighted and the proline repeat in xUBF Box1 is underlined. (**B**) Purification of the mutated Nbox1 protein. Nbox1 and the Nbox1 mutant (Nbox1-mut.) were analysed by SDS gel electrophoresis. The high molecular weight bands in the Nbox1-mut. track were precipitation artifacts.

and/or 3 all played significant cooperative roles in the DNA affinity of xUBF. This was consistent with the protection which HMG boxes 2 and 3 afforded downstream of +20.

The C-terminal segment of xUBF affects the interaction of HMG box1 with the core promoter

Despite the extensive similarity of the Nbox123, Box123 and wild type xUBF footprints, a major difference was apparent (Figure 6B). The hypersites at −11 to −15 within the core promoter, consistently seen with xUBF and referred to as the 'long-range effect' in Figures 2–4, were completely absent in the Nbox123 and Box123 footprints. As was demonstrated in Figures 2 and 3, these were the same hypersites which also required the presence of promoter sequences between +17 and +32. The effect can be clearly seen by comparing the footprints of xUBF and Nbox123 on both the wild type and Δ3′+36 promoters (Figure 6B). Wild type xUBF generated the −11 to −15 hypersites on both DNAs, while Nbox123 was unable to do so on either.

It was concluded that the interaction of HMG box1 with the core promoter was modulated by one or more of the xUBF C-terminal domains deleted in Nbox123, i.e. HMG boxes 4, 5 and/or the acidic domain and that this modulation also required the presence of the ribosomal DNA sequences between +17 and +32. Since HMG boxes 2 and/or 3 interact with the DNA downstream of +22, it is possible that the binding of these domains to the DNA is necessary for the correct positioning of the C-terminal segment of xUBF. Alternatively, HMG boxes 4 and/or 5 might contact the DNA within the +17 to +32 sequence. Since the DNA affinities of xUBF and the Nbox123 mutant were very similar, this explanation, though still possible, must be considered unlikely (see Discussion).

The N-terminal proline repeat subdomain of HMG box1 is implicated in DNA recognition

When we originally investigated the structure of xUBF we identified a semi-conserved proline repeat at the N-terminal of each HMG box (Figure 7A). By homology with the BD peptide of HMGI, the CTD of RNA polymerase II and the SPKK DNA interaction motif (Reeves and Nissen, 1990; Suzuki, 1989, 1990), we suggested that this proline repeat may play a role in the DNA interaction of the HMG box. The proline repeat is found in most HMG boxes of the UBFs as well as of HMG1 and 2 and HMGT (Bachvarov and Moss, 1991). It is not however, evident in the more distant relatives of the HMG box transcription factor family such as SRY and LEF-1 (TCF1α). Despite this lack of sequence conservation, mutations in the exactly analogous region of the HMG box of these factors have more recently been shown to eliminate DNA binding (Nargis *et al.*, 1991; Gething and Sambrook, 1992; Harley *et al.*, 1992). Alignment of the HMG boxes of these proteins does in fact reveal a conserved proline flanked by semi-conserved residues (Figure 7A).

In order to test whether the proline repeat was also implicated in the DNA binding of the xUBF HMG boxes, we introduced point mutations into this region of HMG box1. The two basic residues were changed to acidics, the other changes being semi-conservative. A mutant Nbox1 protein was created using this mutated HMG box1 expressed and purified from *E.coli* as a soluble folded protein (Figure 7B) and footprinted on to the wild type 40S promoter, Nbox1-mutant in Figure 6A. Consistent with a role of the proline repeat in DNA recognition, this mutant showed no detectable binding whatsoever.

Discussion

By studying the DNase I footprinting and gel-shifting of the ribosomal transcription factor xUBF and xUBF deletion mutants on wild type and mutant 40S promoters, it has been possible not only to infer both the DNA and protein moieties implicated but also the general arrangement of DNA and protein within the complex. A single complex was shown to form between xUBF and the 40S core promoter and to

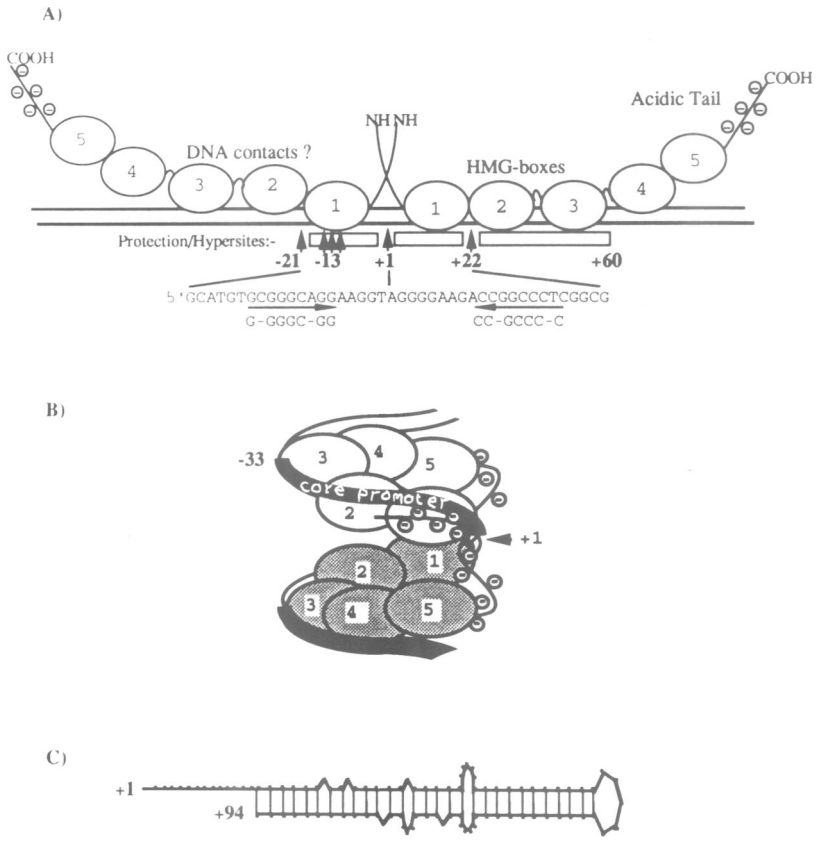

Fig. 8. (**A**) The colinear model of the xUBF–core promoter interaction. xUBF is shown as a dimer with its HMG boxes strung out along the core promoter and 5′ transcribed DNA sequences. A putative DNA recognition sequence is shown below the base positions, which are given relative to the initiation site at +1. (**B**) A hypothetical folded model of the xUBF–core promoter interaction, explaining the manner in which the +17 to +32 DNA sequence and the C-terminal domain of xUBF could affect the interaction of HMG box 1 with the core promoter. A regular super-helical form of DNA is shown purely for convenience, irregular DNA bending also being possible. (**C**) The potential RNA hairpin structure at the immediate 5′ of the 40S transcript as identified and drawn by the UWGCG program FOLD (Devereux *et al.*, 1984).

involve at least the sequences between −22 and +36 and protein–DNA contacts as far downstream as +60. xUBF binding in this region was directed by at least two independent interactions, one around the 40S initiation site, within −22 to +10 and a second downstream of +18, of which a significant portion lay 3′ of +24 and even possibly +32. Footprinting of xUBF deletion mutants indicated that HMG box1 protected two distinct sites, one bounded by hypersites at −20 and −1 and a second by the hypersites at +1 and +22, i.e. on either side of the 40S initiation site. Dimerization via the N-terminal domain of xUBF (McStay *et al.*, 1991a) stabilized this interaction considerably, the K_d decreasing by about two orders of magnitude in the presence of this domain. HMG boxes 2 and/or 3 contacted regions downstream of +22 adding very significantly to the DNA affinity of xUBF (decrease in K_d of ~10 times).

The data suggest a colinear model of xUBF–DNA interaction (Figure 8A). Here for simplicity we have made the likely assumption that a single xUBF dimer is bound within the core promoter and downstream region. As yet no direct DNA binding has been demonstrated for HMG boxes 4 and 5 and our data suggested that these domains did not add to the DNA affinity of xUBF. On the other hand, we showed that the C-terminal segment of xUBF, i.e. HMG boxes 4 and 5 and/or the C-terminal acidic domain (the 'acidic tail'), modified the binding of HMG box1 within the core promoter, such that the region around position −13 became hypersensitive to DNase. This same effect also depended on the presence *in cis* of the distal DNA sequences between +17 and +32, a region in which the HMG boxes 2 and/or 3 contacted the DNA. Thus in the colinear model in Figure 8A, the binding of HMG boxes 2 and 3 downstream of +17 permits the C-terminal segment of xUBF to modify the interaction of the HMG box1 bound between −1 and −20 in the core promoter. Depending on the folding of xUBF, this long-range interaction might just be able to occur on the linear DNA. However, it is likely that a general property of the HMG box domain is to bend or kink the DNA by up to ~130° (Giese *et al.*, 1992; Lilley, 1992). The colinear binding of consecutive HMG boxes of an xUBF dimer to the same face of the DNA should then result in the bending of the core promoter into a short, possibly irregular and probably negative, e.g. see Lilley (1992), superhelix (Figure 8B). Such a structure could bring the C-terminal domain of the downstream xUBF molecule and the HMG box1 of the upstream xUBF in proximity with each other. In Figure 8B this is shown by a purely hypothetical intermolecular interaction between the acidic tail and HMG box 1. Alternative, more complex models in which the interaction between HMG boxes and DNA does not occur in a colinear manner could also explain our data. Such models would nevertheless involve extensive bending of the DNA. We have recently found that wild type xUBF will effectively negatively supercoil plasmid DNA containing the core promoter sequences −9 to +72 (B.Leblanc, unpublished data). This suggests, but does not prove, that

xUBF may be capable of causing significant DNA bending. As yet we have not shown that all the HMG boxes of an xUBF dimer do bind on the same face of the DNA duplex. The binding of HMG box1 between −20 and −1 and +1 and +22 does however, suggest that at least two HMG boxes of an xUBF dimer are positioned on the same face of the duplex (Figure 6A).

In previous studies the C-terminal acidic domain or acidic tail of xUBF and mouse UBF have been shown to be important for transcriptional activation (McStay *et al.*, 1991a; Voit *et al.*, 1992). Our data open the possibility that the acidic tail is also responsible for modifying HMG box 1 binding to the −13 region of the core promoter, making this promoter region easily accessible on the surface of the xUBF−promoter complex. The −13 region of the core promoter has been shown to be crucial for transcription in all species studied, the bases at −7 and −16 being especially important (see Read *et al.*, 1992; Firek *et al.*, 1990 and references therein).

If the downstream xUBF molecule interacts with the upstream one, why does the converse not also occur? The simplest explanation is that it does but that the sequences between +1 and +20 were unsuitable for the DNase to reveal this in the footprint. Alternatively, the proximity of an autonomous interaction of xUBF within the enhancer homology or simply the absence of a good upstream HMG box 2/3 binding site could preclude such an interaction in a manner analogous to the effect of the +17 to +32 and Δ3′+10 deletions (Figures 2 and 3).

Our data suggest that each HMG box of xUBF displays a different DNA sequence preference. HMG box1 bound preferentially to the sequences between −1 and −20 and +1 and +22. Comparison of these sequences revealed an imperfect diad symmetry whose spacing was in accord with the positions of the observed hypersensitive cleavages (Figure 8A). However, DNA sequences downstream of +18 were also found to be sufficient to position xUBF correctly. Hence it must be assumed that the sequence specificity of the xUBF protein does not reside solely in HMG box1 and that boxes 2 and/or 3 also display sequence recognition which is different from that of HMG box1.

In *Xenopus* it has for some time been apparent that xUBF binds around the 40S initiation site (Bell *et al.*, 1989; Read *et al.*, 1992). In mammals the arguments for a similar UBF positioning are indirect but nevertheless convincing. It has been shown that both *Xenopus* and mammalian UBF are essential transcription factors (see Sollner-Webb and Mougey, 1991). Secondly the mammalian core promoter has been especially well characterized as having autonomous promoter activity (see Read *et al.*, 1992 and references therein). Hence, UBF must interact effectively with the core promoter. Like xUBF on the *Xenopus* promoter, hUBF also generates a site of DNase hypersensitivity at −21 on the human ribosomal promoter (Bell *et al.*, 1988). Further, the 3′ boundary of the core promoter in mammals has been mapped to around +16 to +20 and a major interspecies promoter homology runs from +1 to +20 (see Read *et al.*, 1992). The positioning of UBF on the mammalian and *Xenopus* promoters may therefore be similar.

The role of xUBF recognition sites well into the 40S transcribed region is as yet unclear. In *Xenopus* the 5′ of the 40S RNA contains a potential hairpin structure (Figure 8C). This suggested the possibility that xUBF might also bind to this region of the 40S RNA and hence that the transcript might play a regulatory role in its own promotion. A T7 RNA polymerase *in vitro* transcript containing the first 91 bases of the 40S RNA however, failed to compete with the footprint of authentic xUBF on the core promoter, even in large molar excess and under conditions of limiting xUBF (B.Leblanc and T.Moss, unpublished observation). An alternative explanation for the xUBF interaction within the transcribed region could be to prevent complete displacement of xUBF from the core promoter during transcription initiation. DNA strand separation during transcription initation could lead to partial displacement of the xUBF. Interaction of an xUBF dimer both upstream and downstream of the initiation site could prevent permanent displacement of xUBF and still allow the passage of the polymerase. This is essentially equivalent to a previously proposed model for TFIIIA binding on the 5S gene (Miller *et al.*, 1985).

We suggested previously that a proline rich segment at the N-terminal of most HMG boxes might constitute a minor groove DNA binding motif, mainly due to its homology with the BD-peptide of HMGI and the SPKK motif (Bachvarov and Moss, 1991). Subsequently, it was shown that the equivalent segment of SRY and of LEF-1 was essential for DNA binding of these factors and that LEF-1 bound predominantly in the minor groove (Giese *et al.*, 1991, 1992; Harley *et al.*, 1992; Nargis *et al.*, 1991). By mutating this region of xUBF HMG box1, we have now been able to show that it is also important for the DNA binding of xUBF.

Materials and methods

DNA constructions

Most footprinting was performed on the 400 bp *PstI*−*TaqI* (−317 to +92) 40S promoter fragment from pXl108 (Moss *et al.*, 1980; Read *et al.*, 1992) which was either ^{32}P 5′ end-labelled with T4 kinase or 3′ end-labelled using Klenow polymerase at the *TaqI* terminus. The equivalent fragments from the internal deletion mutants Δ−61 to −36, Δ−30 to −13, Δ−18 to +32 and Δ+17 to +32 (Moss, 1982) were also 5′-labelled at the same *TaqI* site. The −35 to +10 and −22 to +10 constructions used in Figure 3B were derived from the Δ3′+10 deletion (see below) using the 'erase-a-base' system (Promega). For footprinting the DNA was 5′ end-labelled at the *ClaI* site of the vector, and then freed from the plasmid by *ApaLI*. Δ5′+24 and Δ5′+32 mutants were also produced with the erase-a-base system. They respectively contain the +24 to +91 and +32 to +91 ribosomal sequences subcloned between the *SstI* and *AccI* sites of pT7T3U19 (Pharmacia). They were 5′ end-labelled at the *AccI* (*TaqI*) site and were then freed from the vector using *BglI*. The Δ3′+36 mutant was a −1050 to +36 ribosomal fragment cloned in the *Bam*HI−*Hin*dIII sites of pBR322 and was 5′ end-labelled at +36. Δ5′−60, Δ5′−9, Δ5′+2, Δ5′+18 mutants were cloned *Bam*HI−*AccI* in pT7T3U19, 5′ end-labelled at +91 (*AccI*/*TaqI*) and freed from the vector using *ClaI*. Δ3′+6 and Δ3′+10 contain ribosomal sequences upstream as far as −245 cloned in the *Bam*HI−*SalI* sites of pT7T3U19. They were 5′ end-labelled at the vector *ClaI* site and then freed from the vector with *PstI*. Δ5′−60, Δ5′−9, Δ5′+2, Δ5′+18, Δ3′+36, Δ3′+6 and Δ3′+10 were subcloned from original deletions kindly provided by B.Sollner-Webb.

In vitro footprinting and gel-shift

Footprinting was performed as previously described (Read *et al.*, 1992). DNA affinities of the recombinant xUBF and xUBF deletion mutants were determined from densitometer scans of footprint analyses executed over a range of known protein concentrations. The 40S initiation site region of each footprint was analysed for the relative degree of protection/hypersensitivity essentially as Brenowitz *et al.* (1991) to yield the dissociation and association constants k_d and k_a. For gel-shifts, various aliquots of xUBF were incubated with 7 fmol of a 5′ end-labelled 155 bp DNA fragment containing the ribosomal promoter sequence extending from −9 to +91 and extended on the 5′ side with plasmid vector (Δ5′−9). The incubation conditions were the same as for the footprinting including 0.5 μg of poly(dA.T), except that the total volume was reduced to 20 μl and 5 mM

$MgCl_2$ was added. The reaction was loaded directly on to a 4% polyacrylamide gel in 25 mM Tris−HCl pH 8.3, 160 mM glycine, 5 mM magnesium acetate, 1 mM EDTA, 2.5% glycerol and 0.5 mM DTT (Ranish *et al.*, 1992). Electrophoresis was performed at constant current (22 mA) at room temperature using the same buffer but lacking glycerol and DTT. Gel shift competition experiments were carried out by adding various amounts of competitor DNA fragment along with the labelled DNA. The specific competitor was simply the unlabelled Δ5′−9 fragment (see above) while the non-specific competitor used in Figure 5 was the *Dde*I−*Hin*dIII fragment from pBR322, bases 4291−29.

In vivo footprinting

Erythrocyte nuclei from fully grown *X.laevis* females were isolated essentially according to the method of Hewish and Burgoyne (1973). Blood was collected in ice-cold 1× SSC (150 mM NaCl, 15 mM sodium citrate) and immediately centrifuged at 1000 r.p.m. (rotor H-6000, Sorvall) at 4°C for 3 min. The cell pellet was washed twice in 1× SSC. The cells were resuspended in 10 pellet volumes of buffer A-0.25: buffer A (60 mM KCl, 15 mM NaCl, 0.5 mM spermidine, 0.15 mM spermine, 15 mM Tris−HCl pH 7.5, 0.1 mM PMSF) containing 0.25 M sucrose, 1 mM EDTA and 0.1 mM EGTA. Triton X-100 (Sigma) and Nonidet P-40 (Calbiochem) were added to 0.1% (v/v). Nuclei were harvested by centrifugation at 2000 r.p.m. and the nuclear pellet was resuspended in buffer A-0.35 (buffer A with 0.35 M sucrose). The nuclei were again recovered by centrifugation and resuspended in 5 pellet volumes of buffer A containing 40% glycerol. 750 μl of sedimented nuclei were added to 3.25 ml of buffer A. The reaction mixture was adjusted to 1.5 mM $MgCl_2$ and 1 mM $CaCl_2$ and DNase I added to 40 ng/μl and digestion allowed to proceed on ice for 30 min with aliquots taken at regular times. These were adjusted to 1% SDS and 10 mM EDTA and phenol−chloroform extracted. Selected aliquots were further digested with *Pst*I, re-extracted with phenol−chloroform and precipitated and 10 μg of DNA from each was subjected to linear PCR amplification (Saluz and Jost, 1989), using the 5′ end-labelled primer 5′CCGGGGACCGAGGCGGGAAACGCCCC3′ (+72 to +47). PCR reactions were performed with 7′-deaza dGTP (Boehringer) in place of dGTP. After 2 min denaturation at 94°C, the reaction was taken through 20 cycles of amplification in an Ericomp thermal cycler: 1 min at 95°C, 1.5 min at 62°C, 3 min at 75°C. The samples were then phenol−chloroform extracted and precipitated, washed repeatedly with 80% ethanol and dried. Electrophoresis was performed on a 0.6 cm thick 6% sequencing gel.

Expression and isolation of xUBF and xUBF mutants

xUBF was prepared from the nuclei of *X.laevis* tissue culture cells as previously described (Read *et al.*, 1992). Vaccinia virus expressed human UBF was a gift from H.-M.Jantzen. Recombinant xUBF2 (the smaller form) was expressed in *E.coli* from the cDNA clone XlUBF2c (Bachvarov and Moss, 1991) inserted at the *Bam*HI site of the pGEX-2T vector (Pharmacia). The recombinant xUBF (amino acids 16−677) was produced as a fusion protein with glutathione-*S*-transferase (GST). Nbox123 was produced by fusing amino acids 16−383 to the GST, Nbox1 amino acids 16−202, Box123 amino acids 101−383 and Box1 amino acids 101−202. Each pGEX construct was transformed into *E.coli* strain HB101 and expressed essentially as described (Smith and Corcoran, 1991). For xUBF, Nbox123 and Box123 the bacteria were lysed by sonication in 20 ml TMc (50 mM Tris−HCl pH 7.9, 12.5 mM $MgCl_2$, 0.1 mM EDTA, 20% glycerol, 1 mM DTT, 0.1 mM PMSF) with 0.07 M KCl and the resulting solution adjusted to 1% Triton X-100. The solution was then centrifuged at 9500 r.p.m. at 4°C for 5 min to remove cellular debris. Further purification followed the previously published procedure for xUBF isolation (Read *et al.*, 1992). NBox1 and Box1 were isolated using glutathione-Sepharose and thrombin cleavage as described (Smith and Corcoran, 1991). Finally these proteins were passed over DEAE-Sephacel (Pharmacia) in 0.15 M KCl, TMc and collected in the flow-through fraction. Protein concentrations of Box1 preparations were determined by UV spectroscopy at 280 nm, using an extinction coefficient of 5100 mol^{-1} cm^{-1} calculated from the predicted amino acid composition. The concentrations of the other constructions were estimated by SDS gel electrophoresis and Coomassie blue staining of various aliquots in comparison with the Box1 preparations.

Mutations of the proline repeat region of HMG box1 in the Nbox1-mut. (Figure 7) were introduced by a PCR-based technique (Landt *et al.*, 1990; Kuipers *et al.*, 1991) using an internal mismatch primer; 5′ GACTGGCTCCTCGAGAAACTCG, where the mismatched positions have been underlined. An additional fortuitous mutation was also reproducibly introduced during the amplification, the G immediately following the 3′ end of the primer being changed to an A. Thus the sequence, PEFPKKPL, of amino acids 109−116, at the N-terminal of HMG box1 was changed to LEFLEEPV. NBox1-mut. was isolated as were the NBox1 and Box1 constructions.

Acknowledgements

We wish to acknowledge the aid of Dr B.Sollner-Webb in providing several promoter deletion mutants and to thank Dr H.-M.Jantzen for providing hUBF, for frank discussions and for sharing unpublished data. We would also like to thank Dr S.I.Dimitrov and Dr A.Ruiz-Carrillo for critical reading of the manuscript, and the latter for as yet unacknowledged critical reading of several other manuscripts. The work was supported by a grant from the Medical Research Council of Canada (MRC) to T.M. and by a collaborative research grant from NATO to C.R. and T.M. (no. 890637). B.L. was supported by a fellowship from the FCAR of Québec. T.M. is a Senior Researcher of the FRSQ and a member of the Centre de Recherche en Cancérologie de l'Université Laval which is supported by the FCAR of Québec.

References

Bachvarov,D. and Moss,T. (1991) *Nucleic Acids Res.*, **19**, 2331−2335.
Bachvarov,D., Normandeau,M. and Moss,T. (1991) *FEBS Lett.*, **288**, 55−59.
Bateman,E. and Paule,M.R. (1986) *Cell*, **47**, 445−450.
Bell,S.P., Learned,R.M., Jantzen,H.M. and Tjian,R. (1988) *Science*, **241**, 1192−1197.
Bell,S.P., Pikaard,C.S., Reeder,R.H. and Tjian,R. (1989) *Cell*, **59**, 489−497.
Brenowitz,M., Senear,D.F. and Kingston,R.E. (1991) In Ausubel,F.M., Brent,R., Kingston,R.E., Moore,D.D., Seidman,J.G., SmithJ.A. and Struhl,K. (eds), *Current Protocols in Molecular Biology*. Greene Publishing Associates and Wiley-Interscience, New York, pp. 12.4.1−12.4.11.
Cary,P.D., Turner,C.H., Mayes,E. and Crane-Robinson,C. (1983) *Eur. J. Biochem.*, **131**, 367−374.
Comai,L., Tanese,N. and Tjian,R. (1992) *Cell*, **68**, 965−976.
Cormack,B.P. and Struhl,K. (1992) *Cell*, **69**, 685−696.
Devereux,J., Haeberli,P. and Smithies,O. (1984) *Nucleic Acids Res.*, **12**, 387−395.
Diffley,J.F.X. and Stillman,B. (1991) *Proc. Natl Acad. Sci. USA*, **88**, 7864−7868.
Firek,S., Read,C., Smith,D.R. and Moss,T. (1990) *Nucleic Acids Res.*, **18**, 105−109.
Gething,M.J. and Sambrook,J. (1992) *Nature*, **355**, 33−45.
Giese,K., Amsterdam,A. and Grosschedl,R. (1991) *Genes Dev.*, **5**, 2567−2578.
Giese,K., Cox,J. and Grosschedl,R. (1992) *Cell*, **69**, 185−195.
Gokal,P.K., Mahajan,P.B. and Thompson,E.A. (1990) *J. Biol. Chem.*, **265**, 16234−16243.
Gubbay,J., Collignon,J., Koopman,P., Capel,B., Economou,A., Münsterberg,A., Vivian,N., Goodfellow,P. and Lovell-Badge,R. (1990) *Nature*, **346**, 245−250.
Guimond,A. and Moss,T. (1992) *Nucleic Acids Res.*, **20**, 3361−3366.
Harley,V.R., Jackson,D.I., Hextall,P.J., Hawkins,J.R., Berkovitz,G.D., Sockanathan,S., Lovell-Badge,R. and Goodfellow,P.N. (1992) *Science*, **255**, 453−456.
Hewish,D.R. and Burgoyne,L.A. (1973) *Biochem. Biophys. Res. Commun.*, **52**, 504−510.
Hisatake,K., Nishimura,T., Maeda,Y., Hanada,K., Song,C.Z. and Muramatsu,M. (1991) *Nucleic Acids Res.*, **19**, 4631−4637.
Jantzen,H.-M., Admon,A., Bell,S.P. and Tjian,R. (1990) *Nature*, **344**, 830−836.
Kolodrubetz,D. and Burgum,A. (1990) *J. Biol. Chem.*, **265**, 3234−3239.
Kuipers,O.P., Boot,H.J. and De Vos,W.M. (1991) *Nucleic Acids Res.*, **19**, 4558.
Landsman,D. and Bustin,M. (1991) *Mol. Cell. Biol.*, **11**, 4483−4489.
Landt,O., Grunert,H.P. and Hahn,U. (1990) *Gene*, **96**, 125−128.
Lilley,D.M.J. (1992) *Nature*, **357**, 282−283.
Mahajan,P.B. and Thompson,E.A. (1990) *J. Biol. Chem.*, **265**, 16225−16233.
Mahajan,P.B., Gokal,P.K. and Thompson,E.A. (1990) *J. Biol. Chem.*, **265**, 16244−16247.
McStay,B., Frazier,M.W. and Reeder,R.H. (1991a) *Genes Dev.*, **5**, 1957−1968.
McStay,B., Hu,C.H., Pikaard,C.S. and Reeder,R.H. (1991b) *EMBO J.*, **10**, 2297−2303.
Miller,J., McLachlan,A.D. and Klug,A. (1985) *EMBO J.*, **4**, 1609−1614.
Moss,T. (1982) *Cell*, **30**, 835−842.
Moss,T., Boseley,P.G. and Birnstiel,M.L. (1980) *Nucleic Acids Res.*, **8**, 467−485.

Moss,T., Mitchelson,K. and De Winter,R.F.J. (1985) *Oxf. Surv. Eukaryot. Genes*, **2**, 207–250.
Nasrin,N., Buggs,C., Kong,X.F., Carnazza,J., Goebl,M. and Alexander-Bridges,M. (1991) *Nature*, **354**, 317–320.
O'Mahony,D.J., Smith,S.D., Xie,W. and Rothblum,L.I. (1992) *Nucleic Acids Res.*, **20**, 1301–1308.
Parisi,M.A. and Clayton,D.A. (1991) *Science*, **252**, 965–969.
Paule,M.R., Bateman,E., Hoffman,L., Iida,C., Imboden,M., Kubaska,W., Kownin,P., Li,H., Lofquist,A., Risi,P., Yang,Q. and Zwick,M. (1991) *Mol. Cell. Biochem.*, **104**, 119–126.
Pikaard,C.S., McStay,B., Schultz,M.C., Bell,S.P. and Reeder,R.H. (1989) *Genes Dev.*, **3**, 1779–1788.
Ranish,J.A., Lane,W.S. and Hahn,S. (1992) *Science*, **255**, 1127–1129.
Read,C., Larose,A.M., Leblanc,B., Bannister,A.J., Firek,S., Smith,D.R. and Moss,T. (1992) *J. Biol. Chem.*, **267**, 10961–10967.
Reeck,G.R., Isackson,P.J. and Teller,D.C. (1982) *Nature*, **300**, 76–78.
Reeves,R. and Nissen,M.S. (1990) *J. Biol. Chem.*, **265**, 8573–8582.
Saluz,H. and Jost,J.-P. (1989) *Proc. Natl Acad. Sci. USA*, **86**, 2602–2606.
Schnapp,A., Clos,J., Hèdelt,W., Schreck,R., Cvekl,A. and Grummt,I. (1990a) *Nucleic Acids Res.*, **18**, 1385–1393.
Schnapp,A., Pfleiderer,C., Rosenbauer,H. and Grummt,I. (1990b) *EMBO J.*, **9**, 2857–2863.
Schultz,M.C., Reeder,R.H. and Hahn,S. (1992) *Cell*, **69**, 697–702.
Sinclair,A.H., Berta,P., Palmer,M.S., Hawkins,J.R., Griffiths,B.L., Smith,M.J., Foster,J.W., Frischauf,A.-M., Lovell-Badge,R. and Goodfellow,P.N. (1990) *Nature*, **346**, 240–244.
Smith,D.B. and Corcoran,L.M. (1991) In Ausubel,F.M., Brent,R., Kingston,R.E., Moore,D.D., Seidman,J.G., Smith,J.A. and Struhl,K. (eds), *Current Protocols in Molecular Biology*. Greene Publishing Associates and Wiley-Interscience, New York, pp. 16.7.1–16.7.8.
Smith,S.D., Oriahi,E., Lowe,D., Yang-Yen,H.-F., O'Mahony,D., Rose,K., Chen,K. and Rothblum,L.I. (1990) *Mol. Cell. Biol.*, **10**, 3105–3116.
Sollner-Webb,B. and Mougey,E.B. (1991) *Trends Biochem. Sci.*, **16**, 58–62.
Stewart,M.A., Hall,L.M.C. and Maden,B.EH. (1983) *Nucleic Acids Res.*, **11**, 629–647.
Suzuki,M. (1989) *EMBO J.*, **8**, 797–804.
Suzuki,M. (1990) *Nature*, **344**, 562–565.
Tanaka,N., Kato,H., Ishikawa,Y., Hisatake,K., Tashiro,K., Kominami,R. and Muramatsu,M. (1990) *J. Biol. Chem.*, **265**, 13836–13842.
Tower,J. and Sollner-Webb,B. (1987) *Cell*, **50**, 873–883.
Travis,A., Amsterdam,A., Belanger,C. and Grosschedl,R. (1991) *Genes Dev.*, **5**, 880–894.
Voit,R., Schnapp,A., Kuhn,A., Rosenbauer,H., Hirschmann,P., Stunnenberg,H.G. and Grummt,I. (1992) *EMBO J.*, **11**, 2211–2218.
White,R.J., Jackson,S.P. and Rigby,P.W.J. (1992) *Proc. Natl Acad. Sci. USA*, **89**, 1949–1953.
van de Wetering,M.L., Oosterwegel,M.A., Dooijes,D. and Clevers,H.C. (1991) *EMBO J.*, **10**, 123–132.

Received on September 9, 1992; revised on November 16, 1992